DOCUMENT RESUME

ED 334 238 TM 016 765

AUTHOR Ejzenberg, Roseli
TITLE Oral Fluency in the Context of Second/Foreign Language Testing: Understanding the Role of Task Structure and Formulaic Speech.
PUB DATE 90
NOTE 33p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adult Education; Adults; Data Collection; English (Second Language); Foreign Countries; Higher Education; *Language Fluency; *Language Proficiency; *Language Tests; Models; Postsecondary Education; Rating Scales; Research Methodology; *Second Language Learning; *Speech Skills; Test Construction
IDENTIFIERS Brazil; *Formulaic Speech

ABSTRACT

A study was planned to generate a model for testing second language education that would incorporate task types to allow the display of fluency. It is argued that the display of oral fluency can be used as a selection criterion for inclusion of test items in multitask or portfolio assessment. The following three research questions were formulated: (1) how different speaking tasks varying in structure in terms of cuing and expected response affect the display of oral fluency; (2) the difference in oral discourse produced in the performance of the four research tasks; and (3) the task-specific nature of production of formulaic speech. Young adult students (N=32) from the highest English fluency levels of a school for English in Brazil would produce cued and uncued dialogues and picture-cued or uncued narratives (monologues). Testing sessions were to be evaluated by two examiners. It was expected that: task structuring involving manipulation of cued/uncued and dialogue/monologue conditions would affect the amount of formulaic speech produced, and the fluency ratings obtained on the tasks with cued tasks would generate more formulaic and more fluent speech. Data to be analyzed include: 128 audiotaped speech samples from the subjects; subjects' proficiency and fluency scores on tasks; subjects' written responses to a survey questionnaire; and subjects' responses to semi-structured interviews. The test conditions and documents to be used in the study are appended. A 74-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED334238

ORAL FLUENCY IN THE CONTEXT OF SECOND/FOREIGN LANGUAGE TESTING:

Understanding the role of task structure and formulaic speech

Roseli Ejzenberg
SUNY-Albany

Outline of the Paper

1. Overview of the Problem
2. Literature review on Native and Nonnative Oral Fluency
   - 2.1 Fluency-Related Phenomena
   - 2.2 Variability in Second/Foreign Language Production
     - 2.2.1 Effect of Task on Oral Production
     - 2.2.2 Effect of Task on Oral Testing: Two Approaches to Oral Proficiency Assessment
3. Conceptual Framework
   - 3.1 The fluency construct: variables under study
   - 3.2 Research Questions
4. Study Design:
   - 4.1 Research Site and Sampling
   - 4.2 Task Design: (See Appendix A for "instruments")
   - 4.3 Task Sequencing (See Table 1)
   - 4.4 Data Collection
   - 4.5 Data Analysis
     - 4.5.1 Statistical Analysis
     - 4.5.2 Discourse Analysis
   - 4.6 Anticipated Outcomes
5. References
6. Appendices

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
ROSELI EJZENBERG
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM016745

BEST COPY AVAILABLE

ORAL FLUENCY IN THE CONTEXT OF SECOND/FOREIGN LANGUAGE TESTING: UNDERSTANDING THE ROLE OF TASK STRUCTURE AND FORMULAIC SPEECH

Roseli Ejzenberg

"Just as a wave is not defined by the water it moves through, fluency is not defined by the sum of its constituents. It 'appears', as if by magic, when all the conditions are right." (Riggenbach, 1989, p. 174).

1.Overview of the Problem

Much of what happens in foreign language education and testing in New York State these days is affected by the Foreign Language Proficiency an Regents Exams. The heightened attention testing has received since the 80's transcends the local and the state level, though. Emerging amidst what has been considered a nationwide "revolution" in language testing (Clark, 1986, p. 131), the proposed study is an attempt to address some of the pending problems: 1. How should we test speakers at higher levels of second/foreign language proficiency?; 2. What are we really testing at these levels: proficiency, fluency or strategic competence (Canale & Swain, 1980)? How are these three related?; and a more fundamental question: 3. What are fluency and proficiency in a second/foreign language?

After several years of debate on whether language proficiency is a unitary trait or a composite of various component abilities (See, for example, Oller, 1979; and Carroll, 1983), language ability is generally conceived as being formed by (...) "several distinct but related constructs in addition to a general construct of language proficiency". (Cf Oller, 1983a in Bachman, 1990, p. 82)

Surprisingly, the systematic endeavor to incorporate a theoretical framework with the methods and technology for measuring language proficiency -- more recently called "communicative language ability" (Bachman, 1990) -- is fairly new in the testing field. There have been very few attempts to provide an operational definition of Language Proficiency and establish its construct validity, so currently much research on the nature of communicative language ability has focused on identifying and empirically validating its components. Most attempts to validate such components, however, have been inconclusive. (ibid., p. 86).

Despite the "crucial" role fluency plays as a major component of foreign language proficiency (Graham, 1978; Mullen, 1978, and Oller, 1979), its relationship with it is still unclear. (Butler-Wall, 1986; Hieke, 1985). Little research has applied findings from other research disciplines to the second/foreign language context; few empirical studies have examined the nonnative fluency construct (Butler-Wall, 1986), and there is little understanding of what exactly comprises fluency. (Riggenbach, 1989).

Fluency is a component of language proficiency that is present with different degrees of salience across proficiency levels becoming

more salient as a speaker moves towards higher proficiency levels. Whereas a novice or a less proficient speaker is still "groping for words" (Foreign Service institute fluency scale cited in Vallette, 1979, p.) and struggling to get a message across, a more advanced speaker will be concerned about how fluently that message will be conveyed. This is clearly instantiated by the well known ACTFL (American Council for the Teaching of Foreign Languages) Proficiency Scales in which fluency goes from not even being mentioned in the lower proficiency levels to being directly included in the description of the higher levels. (Riggenbach, 1989, pp. 154-156). This poses the question of whether fluency is a "mysterious quality that appears at a certain level of oral proficiency". (p. 156)

Specific understandings of what fluency is affect both the development and design of curricular, instructional and testing materials as well as the corresponding educational praxis. For instance, "oral proficiency tests often assume the rater's ability to make intuitive judgements about fluency, describing it as 'ease of communication' or 'speech which lacks unnatural pauses'. It is not clear whether such subjective criteria reflect the existence or lack of any concrete features in the language itself." (Riggenbach, 1989, p. xiv).

The main purpose of this study is to generate the underpinnings of a model for testing second language which incorporates task-types that allow for display of fluency, (i.e., tasks with potential to yield performance variability at higher levels of proficiency). It is my contention that potential for display of Oral Fluency can be used as a key selection criterion for inclusion of test items in Multitask or Portfolio Assessment when testing speakers at higher levels of Oral Proficiency.

2. Literature Review on Native and Nonnative Oral Fluency

Native and nonnative oral fluency -- variously described as: "speech which is effortless and smooth. (SPEAK Test rating scale)." (ibid, p. 1), "a quality of speech that is produced in novel situations" (Yorozuya, 1982), or "ability to plan and arrange (...) speech performance (and) to avoid numerous very long pauses." (Cf. Raupach, 1980, p. 268 in Butler-Wall, 1986, p. 323) -- has been the object of study in several sub-fields of research.

In the field of Education multidisciplinary studies drawing on Applied Linguistics, Sociolinguistics and Pragmatics have focused on Discourse and Interactional analysis as tools to better understand fluency and fluency-related phenomena (such as "disfluencies", i.e., pause, hesitation and repair behavior) in both first and second/foreign language. (Butler-Wall, 1986; Riggenbach, 1989). Within this strand of research one theoretical framework, called a "sociocommunicative/sociocognitive view of language and literacy" (Langer 1987), has guided the discussion of the relationship between fluency and language proficiency. Both are studied as they affect and are affected by a variety of cognitive, social and cultural-specific systems and practices. (ibid).

2.1 Fluency-Related Phenomena

In order to understand the notion of fluency we need to examine its components (i.e., its constituent parts) and its features (i.e., its qualities or properties). In trying to synthesize the most salient fluency-related phenomena in the literature I encountered two main problems. First, definitions of fluency in the existing studies vary considerably in the scope of the phenomenon described. Fluency has been described at different levels of occurrence, namely: at the utterance (E.g: Hieke, 1988) versus at the discourse level (E.g: Goffman, 1981; Jefferson, 1979), at the phonetic (E.g: Hieke, 1984a; Jull, 1987) and grammatical (E.g: Murphy, 1986) versus at the semantic (i.e., information flow such as in Hieke, 1984b and Schaefer, 1980) and the sociolinguistic or functional-interactional level (Kramsch, 1986). It is posited that this confusing state of affairs is symptomatic of a basic misconception in the way fluency-related phenomena have been addressed in many studies, namely, the lack of a situational framework of analysis (Cathcart, 1986), which examines speech production within its context of occurrence.

The second problem emerged when trying to distinguish between the different components and features of fluency because they overlap. For instance, "utterance length" and "complexity" are more static components of fluency determined among other factors by number of words in the clause, phrase or utterance. These components of fluency contribute to dynamic features of fluency, such as "speech rate" and "flow or continuity", which are temporal variables centered around the notion of amount of language over time.

Besides, there is a danger of looking at components of fluency in isolation and apart from the dynamics of language use since we may gain some insight into "what contributes (sic) to fluency" instead of "what fluency itself is". (Riggenbach, 1989, p. 163). It was decided then that in the foregoing discussion, both components and features will be indiscriminately called fluency-related phenomena.

The following fluency-related phenomena appear in the literature: rate of articulation, linking, pause and hesitation phenomena (called disfluencies), repair phenomena, flow of speech or continuity, utterance length and complexity , and accessibility of language (also called ease of expression). Unfortunately, time and space constraints do not allow us to discuss these in this much shortened version of my proposal, so we will zero in on the feature most pertinent to the study.

Underlying the notion of fluency we find its counterpart "disfluency" for in order to understand the notion of continuity, for instance, we need to look at the frequency of disfluencies in speech. Disfluency markers include hesitation and repair phenomena such as filled and unfilled pauses, repetitions, restarts and fillers. The phenomenon of pausing is central to the notion of fluency. Extensive experimental data show that (...) "pause or slow down at or near clause boundary is the normal pattern in lengthy connected discourse, even among speakers who are heard as very fluent". (Pawley & Syder, 1983, p. 200). On the other hand, Goldman-Eisler's studies (cited in

ibid., p. 222) show that "clause-internal pauses are most frequent when the cognitive task is most difficult, as in a first formulation of ideas about a newly observed set of relationships (e.g. in stating the point of a cartoon), and that when the task is easier, hesitations at places other than major grammatical boundaries diminish markedly."

Solid evidence (Cf. Pawley & Syder, 1975, 1976; Chafe, 1979 cited in Pawley and Syder (1983, pp. 202) shows that the longest unit of discourse that can be "fully encoded in one encoding operation" is a single clause containing from 8 to 10 words. This "one clause at a time constraint" on the planning of novel speech is compensated by two types of strategies: (a) a "clause-integrating strategy", i.e., to commit to grammatical constructions which require the speaker to consider the syntax of previous or subsequent clauses while formulating a new one or (b) a "clause chaining strategy", i.e., to "string together" a series of clauses which are more independent from previous or subsequent ones. (ibid.)

The authors point out (pp. 203-204) that while more mid-clause pauses and structural breakdowns correlate with use of the clause-integrating (syntactic) strategy, the chaining strategy reduces the possibility of disfluency or breakdown. There are two advantages in using the chaining style. First, since each major semantic unit is only a single clause, it can be encoded and uttered without internal breaks (p. 203), and second, because the speaker can maintain grammatical and semantic continuity by juxtaposing clauses planned more or less independently to form "multi-clause fluent units". These stretches of "pause-free, promptly delivered speech extending over two or more clauses" (p. 204) correspond to what has appeared in the literature as "holistically stored sequences" or "holophrases" (Corder, 1973), "prefabricated routines and patterns" (Hakuta, 1974, Krashen & Scarcella, 1978), "formulaic speech" (Wong-Fillmore, 1976), "gestalt language" (Peters, 1977) and "language exponents" (Van Ek, 1975). These chunks of language vary in length and range from fixed to semi-fixed structure, being stored in memory and retrieved as "preassembled chunks" (Cf. Bolinger, 1975) or units as much as a single word. (Nattinger, 1985, p. 3).

Prefabricated speech has the advantage of more efficient retrieval and of letting speakers focus on the larger structure of discourse rather than on individual words. Furthermore, reliance on prefabricated speech accounts for has been called "Nativelike Fluency": "the ability of native speakers to produce fluent stretches of spontaneous connected discourse" (although) "human capacities for encoding novel speech in advance or while speaking appear to be severely limited." (Pawley & Syder, 1983, p. 191).

2.1 Variability in Second/Foreign Language Production

There are four sources of variability in interlanguage production: (a) the speaker's idiosyncratic processing capabilities; (b) the second language production itself viewed as a "juggling act" in which attention to one area (e.g.: accuracy) may impair another (e.g; speech fluency). Thus, "systematicity at one level may leave variability at another" ; (c) the analytical or rating criteria that

may affect the perception of variability in a speaker's language production, and (d) the situation(s) or context(s) of speech production, which have to do with task, setting and participants. (Schachter, 1986a, p. 119).

A full account of fluency requires an analysis of its occurrence at both the utterance and the discourse level because fluent performance at the utterance level may be disfluent at the discourse level. (second source of variability). See, for example, Goffman, 1981; Jefferson, 1979 or Butler-Wall, 1986, p. 284 for a discussion of the adverse effects of "relentless" or hiper fluency.

Successful speaking is actualized through linguistic (i.e., grammatical), discoursal, sociolinguistic and strategic competencies (Canale & Swain, 1980) literally "co-operating" -- that is, operating in conjunction and helping each other. These competencies are differentially called into play depending on the context of speech production (the fourth source of variability) in a broad sense and specifically depending of the speaking task at hand.

Whereas a speaker may be very successful at the utterance level, when given a task that requires production of extended discourse s/he may be at a loss. (first source of variability). The different demands on speech production in short and long turns, i.e., at the utterance versus at the discourse level, is cogently presented by Nunan (1989, pp. 27-28) citing Brown and Yule (1983, pp 19-20).

> "The ability to give an uninterrupted oral presentation is quite distinct from interacting with one or more other speaker for transactional and interactional purposes (...) (specially) in short, interactional exchanges in which the learner is only required to make one or two utterances at a time."

In the field of testing there seems to be less of such clear awareness about the distinction of fluency at the utterance and at the discourse level. This is revealed upon examination of the rating scales of testing instruments such as the as the ACTFL Oral Interview. On these tests fluency is treated as one among a number of discrete skill areas such as vocabulary, grammar, and pronunciation. "Such rating scales seem to be addressing primarily the utterance level." (Riggenbach, 1989, pp. 10-11). (Cf. Oller, 1979, Savignon, 1985, Higgs & Clifford, 1982, and Bachman & Savignon, 1986 for discussion of these rating scales).

2.2.1 Effect of Task Type on Oral Production

We saw in the previous section that once source of variability in speech production is its context of occurrence or use. "Any act of speech is functionally organized (that is, it is as attempt to **do** something) for a particular situation in relation to a particular topic." (Finocchiaro & Brumfit, 1983, p. 13). Thus, a speech act (Austin, 1962) is characterized by the need to perform a communicative function such as making an invitation or asking for information. (Omaggio, 1986) in a given context which implies both the sociolinguistic and cultural situation as well as the particular

speaking task that serves as the micro-context of language production.

An extensive body of research on the effects of task on Oral language production points to differential language yield as a function of task type. (Long, 1980, 1981 and 1983; Doughty & Pica, 1986; Porter, 1983; Young, 1984; Duff, 1986; and others). "One way to conceive of a language test is as a means for controlling the context in which language performance takes place. From this perspective, the characteristics of the test method can be seen as analogous to the features that characterize the context of situation, or the speech event, as this has been described by linguists" such as Hymes (1972) and Labov (1972). (Bachman, 1990, p. 111). "The particular way test method facets are designed and controlled, and the correspondence between these and the features of language contexts will have a direct effect on the 'authenticity' of the test and test tasks." (p. 112)

2.2.2 Approaches to Oral Proficiency Assessment

Two main approaches have emerged as a result of two distinct definitions of oral language ability and the importance of task authenticity (i.e., perceived resemblance to real life language use) in testing: the Proficiency approach (See, for example, Higgs & Clifford, 1982; Lowe, 1988) and the Interactional/Ability approach (Bachman, 1990, p.5). The first, also called the "Real Life" approach, (ibid., p. 41) posits that "(...) content and context must be built into the definition of proficiency." (Higgs & Clifford, 1982, p. 60). This approach identifies domains of actual language taking into account (...) "virtually all the features that are present in any instance of language use" (Bachman, 1990, p. 41), striving to create more authentic testing situations.

This first approach had a boom especially in the 80's as a "revolution" (Clark, 1986, p. 131) which came to be known as the "Proficiency Movement". (See Higgs, 1982 and 1984; James, 1985 and Ervin, 1983). Its hallmarks are the American Council on the Teaching of Foreign Languages (ACTFL) "Proficiency Guidelines" and the "ACTFL Oral Interview", (Bachman & Savignon, 1986, p. 380), which followed the Interagency Language Roundtable (ILR) and the Foreign Service Institute (FSI) oral interviews.

Critics of the first approach claim that the ILR and ACTFL scales are (...) "highly dependent upon context factors" (Bachman & Savignon, 1986, p. 386) and, consequently, the ratings are restricted to the particular situations for which they were designed. The problem is that the more narrowly communicative language proficiency is defined with respect to forms, functions and content, the less generalizable a measure it will be. (pp. 386-387). Advocates of the alternative approach, the "Interactional/Ability approach" (Bachman, 1990, p. 42) define language in terms of its "component abilities" such as Canale and Swain's (1983) four levels of communicative proficiency, namely: linguistic, discourse, sociolinguistic and strategic competencies. They aim at testing "pragmatic competence" (...) "the knowledge necessary, in addition to organizational competence, for appropriately producing or comprehending discourse. Specifically, it includes illocutionary competence, or the knowledge

of how to perform speech acts (...)." (Bachman, 1990, p.42)

One of the purposes of the present study is to investigate whether the strong face validity of conversational tasks such as role-play and oral proficiency Interviews (advocated by the first approach) might have been attributed somewhat unwarranted validity in the field of testing. According to Underhill (1987, p. 56) these types of task are an inadequate test for discriminating effectively at higher levels of proficiency. This may be accounted for by the fact that they do not allow for optimal display of proficiency at the discourse level nor do they allow for a display of extended discourse "fluency".

3.CONCEPTUAL FRAMEWORK:

3. A preliminary understanding of the Oral fluency construct:

A selective synthesis of the main research findings about the fluency construct as they pertain to the specific variables under study is presented below:

1. The fluency subscale was "the clearest predictor of overall proficiency as measured on this FSI interview, and thus could be the most reliable of the subscales." (Cf. Hinofotis, Bailey & Stern, 1981 cited in Riggenbach, 1989, p. 157). In other words, fluency is a potentially strong indicator of oral proficiency, specially when testing more proficient speakers. **9**

2. "Fluency has been shown to be sensitive to context" (Cf. Ochs, 1979 in Olynyk et al, 1983, p. 216) and to be differentially displayed accordingly. (Riggenbach, 1989). Olynyk et al (1983, p. 216), for instance, examined fluency in planned versus unplanned discourse and Brown and Yule (1983), who studied "communicative stress" (p. 34), found that a speaker will do best, i.e., perform most fluently, under conditions of least "communicative stress". Three conditions contribute to higher or lower degrees of "communicative stress": a) "features of the context"; b)"state of knowledge of the listener", and "type of task, including status of knowledge and task structure. Task structure, operationalized as "cued/uncued condition" and "dialogue/monologue condition", is the variable under consideration in the proposed study.

3. Cuing, also called "external support" or "structure" to a task (Brown & Yule, 1983, P. 36), is an important variable in determining task structure so that finding the appropriate balance in the amount of cuing inserted in the input will determine the speech outcome of the testing situation. The effect of task cuing is exemplified below:

> (...) "It is helpful to the speaker if the information in the task provides its own structure so that the language is externally supported by the requirements of the task. Thus it is easier for any speaker to give an account of a series of events than it is to provide an argument for 'why' those events

occurred in that order." (ibid., p. 34)

Too much cuing will stiffle production and prevent the examinee from showing autonomous language production. Besides, it is desirable to deliberately insert tasks with increased communicative stress when testing more proficient speakers by, for instance, asking an examinee (...) "to create a structured account of an event which has no obvious structure." (Brown & Yule, 1983, p. 36)

Yet, tasks that fall short of supplying enough cuing can also bias the testing performance so that the examinee may be mistakenly perceived to be less fluent than s/he really is:

> "Some tasks are so open that the speaker does not know what constitutes a good answer, so hesitation and brief replies cannot be taken as evidence that the speaker can't use the foreign language." (p.35)

4.The amount of "communicative stress" (ibid) present in a task that demands display of fluency in "short turns", i.e., (...) 'only one or two utterances" is much less than in a task that demands display of fluency in "long turns", i.e., (...) "a string of utterances (...)". (Brown & Yule, 1983, p. 16). This distinction refers to what has discussed as expected speech response at the utterance level and at the discourse level. Clearly, (...) "short turns do not demand much of the speaker in the way of producing structure." (p. 16). On the other hand, "as soon as a speaker 'takes the floor' for a long turn, tells an anecdote, tells a joke, explains how something works, justifies a position, describes an individual, and so on, he takes responsibility for creating a structured sequence of utterances which must help the listener to create a 'coherent' mental representation of what he is trying to say." (p. 17)

5.. Fluent speech is characterized by greater engagement in what has been called use of a chaining strategy as opposed to a grammmatical integration strategy. (Pawley & Syder, 1983)."Most spoken language consists of paratactic (unsubordinated) phrases which are marked as related to each other, not so much by the syntax as by the way the speaker says them. The speaker uses the resources of pausing and rhythm, and, to a lesser extent, intonation, to mark out for the listener which parts of his speech need to be co-interpreted. Of course what syntax there is will contribute to this structuring, but it is frequently the case that the syntax is rather simple." (Brown & Yule, 1983, p. 4 based on actual speech samples of NSs).

6. "Fluent speech contains many fixed expressions and prefabricated chunks." (Olynyk et al, 1990, p. 144; corroborated by Bygate, 1987; Pawley & Syder, 1983)

7. "An increase in frequency of occurrence of hesitation phenomena correlated with the creation of novel utterances or uncertainty." (ibid.) Thus, speech rate increases when a speaker is producing formulaic speech. (**get source from Jill Kinkade**)

8. "Languages differ in the patterning of hesitation phenomena"

(ibid) and fluent speech. Although fluency and pause profiles are idiosyncratic, they do vary sistematically across linguistic communities (E.G:Scollon & Scollon, 1981 and 1983) and even within the same individual across languages. (Olynyk et al, 1990, p. 141). The present study will be conducted with a homogeneous linguistic group whose native language is Portuguese, and systematic variability in their nonnative English speech production is expected.

3.2. **Research Questions**

It is clear that the much needed reconceptualization of the fluency construct needs to occur within a Situational framework (Cathcart, 1986). Despite the extensive research on the effects of task type on oral production and despite the instrumental role played by formulaic speech in language proficiency and fluency, there is very little research analyzing the influence of task on fluency and on the production of formulaic speech. Although several studies show that methods used to measure language ability influence performance on language tests, I have found no studies to date attempting to investigate the effect of foreign language testing tasks on perceptions of examinee's oral fluency as this may be related to formulaic speech production.

Only by examining the effect of context on discourse production as emergent from task type, participant roles and speakers' perceptions of task demands will we be able to start establishing the underpinnings of a model which incorporates situational variables to understand how fluency is differentially displayed across task-types and settings.

Once we obtain a model of variable fluent speech production across tasks, then we may control for task-effect in order to examine the fluency-internal features that comprise fluency. Fluency-internal features are here understood as those more strictly derived from the domains of linguistic and discourse competence. One of these features is formulaic speech which seems to be the missing link to solve the puzzle of nativelike fluency. (Pawley & Syder, 1983).

Profiles of very fluent and very nonfluent speakers (Riggenbach, 1989, p. 174) should then be interfaced with task-types to fully understand the effect of task demands on display of fluency as mediated by the speakers' perceptions of such demands. (Samuda, 1990).

Three research questions have been formulated:

A. How do different speaking tasks varying in task structuring (test method) with respect to cuing and expected response affect the display of oral fluency? Do an examinee's relative or partial fluency ratings differ across tasks?

B. Is there a difference in oral discourse produced in the performance of the four tasks? If so, is it related to variation in a speaker's ratings across tasks?

C. A related question is: Does production of formulaic speech vary

across tasks? In other words, is it task-specific? If so, is a speaker's perceived fluency related to variability in formulaic speech production across tasks?

4. STUDY DESIGN

4.1. Research Site and Sampling:

Private English as a Foreign Language (EFL) School located in Sao Paulo, Brazil. The school caters to liberal professionals and college students learning English for professional or academic advancement, future travel or as a hobby. The population is of young adults ranging from 18 to 35 years old. The number of students and groups (classes) within each level bracket changes each term according to enrollment. According to demographics from the spring semester there were 400 students in the school 45 of whom were in the highest level bracket, but with promotion of students from lower levels and expected increase in enrollment in the fall semester the population size will increase. 32 students will be randomly picked from the highest EFL level bracket in the school. The sample will be stratified (Brown, 1988, p. 112) according to gender if fall demographics permit.

This site and population were chosen for two reasons. First, this EFL school offers a very wide range of proficiency levels. (Altogether, there are 12 levels clustered into 6 level brackets.) The second reason is that given the exploratory nature of the study and the complexity of the variables already built in, it seemed desirable to control for developmental factors that would have been inserted had a non-adult group of second/foreign language learners been chosen.

4.2 TASK DESIGN

Design: Compare fluency ratings of students' discourse production on two types of role-play with examiner (cued and uncued dialogue) and two types of narrative (picture-cued versus uncued monologue). These tasks will vary in degrees of task structuring or "test method facet" (Bachman, 1990, p. 119) along two variables that have been found to affect discourse production, namely: "task input" (ibid.) here operationalized as cued or uncued condition and "expected response" (ibid.) here operationalized as dialogue and monologue condition or genre.

As it can be seen in the task prompts (See Appendix A), the difference between the cued and uncued conditions in the two dialogue tasks will be respectively operationalized as presence and absence of turn-by-turn cues or directions on the speech act to be performed in the role-play with the examiner. In the monologue tasks the cued and uncued conditions will be respectively operationalized as presence and absence of external support for the narrative. Thus, in the picture-cued monologue, the speaker is relieved from the burden of searching for a way of organizing the narrative since the organization is imposed by the sequencing of the pictures themselves. In the uncued monologue, on the other hand, the examinee will have to make decisions

about the rhetorical organization of the narrative.

It is important to note that the tasks included in the study entail primarily transactional speech. (Brown & Yule, 1983, p. 11) Four illocutionary acts have been selected, namely: requesting a service, imparting anecdotal/personal information, narrating a sequence of past events, and persuading. These four are expected to be perceived as part of the realm of communicative experience of the examinees.

4.3 Task Sequencing:

Examinees will be asked to perform four tasks in counterbalanced order in order to control for "response set". (Harlow, 1990, pp. 229-330). Task sequencing definitely has an impact so that examinees will probably do better on the third and fourth tasks even if the tasks ask them to perform different illocutionary acts dealing with different topics. Consequently, it was decided to sequence tasks so that there would be one monologue and one dialogue in each half of the test (i.e., first half= 1st and 2nd tasks to be performed; second half= 3rd and 4th tasks).

Furthermore, there is a learning effect that takes place from a cued to an uncued condition in the dialogue so that if an examinee performs the cued dialogue first s/he is likely to transfer the dialogue structuring as instantiated by turn taking pattern, expected length of response and interaction to the uncued dialogue. (SEE TABLE 1). The solution was to counterbalance the task sequence so that in half of the tasks (N=16) the cued dialogue (CD) will precede and in the other half (N=16) the uncued dialogue (UD) would precede. These two conditions are respectively called CD-first test and UD-first test in Table 1.

As seen in the table, the counterbalancing has yielded 16 possible task sequences or tests. Since the Total N is 32 examinees (i.e., 128 speech samples) each test combination (i.e., task sequence) will be administered to two randomly assigned examinees (if possible to a male and a female according to our stratified sampling)

Table 1- **Task Sequencing= SEQ**

TASK SEQ. No.	"UD-FIRST" tests								"CD-FIRST" tests							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1st	UM	UD	UM	UD	CM	UD	CM	UD	UM	CD	UM	CD	CM	CD	CM	CD
2nd	UD	UM	UD	UM	UD	CM	UD	CM	CD	UM	CD	UM	CD	CM	CD	CM
3rd	CM	CM	CD	CD	UM	UM	CD	CD	CM	CM	UD	UD	UM	UM	UD	UD
4th	CD	CD	CM	CM	CD	CD	UM	UM	UD	UD	CM	CM	UD	UD	UM	UM
	Subtotal: 8 combinations								8 combinations							

Total: 16 possible combinations

Abbreviations: cued dialogue = CD cued monologue = CM

uncued dialogue = UD uncued monologue= UM

Another possibility for task sequencing would be to go one stap beyond the counterbalancing procedure described above (which yielded 16 task combinations or tests) and formulate a second hypothesis about sequencing. This would further reduce the number of possible task sequences as it can be seen in Appendix E. Given the small sample size in the present study, it was decided not to follow the latter possibility.

4.4 DATA COLLECTION

Data will include (a) 128 audiotaped speech samples from 32 randomly selected students' performance on the four tasks, (b) proficiency and fluency scores on tasks as rated by the three raters who will listen to randomly ordered segments of taped speech samples, (c) written survey questionnaire and (d) semi-structured interviews with 2 examiners, 3 raters and 6 examinees chosen as case studies. First Phase: Audiotaping speaking tasks

Performance of four selected oral tasks will be tape-recorded on the site outside of the classroom. The 35-minute testing sessions will be conducted by two trained examiners, chosen among the senior teachers. The testing sessions will be held outside of class within the context of Diagnostic Placement Exams in the beginning of the school term.

The testing session will start with a three-minute warm up in which examinees will be asked some nonthreatening personal questions in order to relax and feel more at ease with the examiner. The latter

will also provide a brief explanation of the testing procedure. The test will consist of 4 different tasks, and instructions for each task will be given prior to performing it. Each task is expected to last a maximum of 5 minutes without counting instruction time.

English will be spoken throughout and all tasks will be presented in English. However, examinees will be asked whether they understand instructions prior to performing each task and whether they need any further explanation. If they do, the examiner will offer an explanation first in English and then in Portuguese if necessary.

Second Phase: Rater Training

Two rater training sessions will be conducted with the three raters using a sub-sample of 24 tasks (6 students performing 4 tasks each) will be randomly derived from the original corpus of 128 tasks (32 students performing 4 tasks each). Then the 24 recorded tasks from the sub-sample (of 6 students) will be coded and randomly scrambled to avoid "back-to-back" scoring of tasks performed by the same examinee. (Clark & Swinton, 1979). Finally, two tapes containing each 12 randomly picked tasks incognito and in scrambled order will be made.

The first tape will be played to 3 EFL teachers from the school who will be asked to give a holistic rating of the students' fluency on each task using a 6-point scale. The tape will be stopped after each task is played and raters will have one-two minutes to write any comments they might have on the examinee's fluency. At the end of the tape raters will be asked to talk about their ratings and make explicit the criteria that have guided their ratings. Finally, they will be asked to reach a consensus on how fluency should be rated.

In order to test interrater reliability the second tape containing 12 tasks incognito and in scrambled order will be played and rated. Lastly, ratings will be compared and a final procedure for the rating of the speaking task sample to be used in the next phase of the study. This sample will contain 104 speech samples resulting from 26 examinees' performance of four tasks each.

Third Phase: Fluency Rating sessions

There will be six rating sessions. Each session will be divided in two one-hour-long sittings with a 15 minute break each. In each session the three raters will be asked to rate the students' fluency on approximately 20 tasks (i.e., a sixth of the corpus) on a 6-point holistic scale ranging from "very fluent" to "not fluent at all". One minute will be allowed at the end of each rating for raters not to be influenced by the last task and for them to jot down any comments/justifications for their ratings. The comments will be optional for ratings ranging from 2 to 5 and required for ratings on the extremes of the scale (i.e. 1 and 6). For example: Subject 18/ Task 2/Fluency Rating "1"/ Comment: "very thick accent gets in the way/ extreme difficulty pronouncing certain sounds makes examinee hesitate all the time." (See Appendix D for Fluency Rating Form)

Fourth Phase: Scoring

The ratings assigned to each task by the three raters will be mathematically averaged to form each examinee partial fluency rating on each task. Then, each examinee's four partial fluency ratings (for each task s/he will have one average of the three raters' ratings) will be added and mathematically averaged into one Overall Fluency Rating per examinee. Finally, examinees will be rank ordered with respect to their overall fluency ratings in a distribution going from highest to lowest.

Fifth Phase: Case Studies of High and Low fluency examinees

It is anticipated that examinees will vary in their (...) "ability to manage certain aspects of discourse, making some sound more fluent than others." (Olynyk et al, 1983, p. 216). Following the recommendations for further research resulting from the Riggenbach (1989) and Olynyk et al (1983) in phase five of my study profiles of high and low fluency speakers will be interfaced with task types.

Thus, the fifth phase will consist of discourse analysis of speech samples from 6 selected case studies. The cases studies will be selected from the top and bottom 10th percentile of the rank ordered distribution, namely the three examinees with the highest and the three examinees with the lowest overall fluency ratings. Finally, with due regard for confidentiality, the tape recorded data from the 6 selected case studies will be transcribed and subjected to discourse analysis by the researcher.

4.5. Data Analysis

Data analysis procedures will be divided in two phases: in the first phase statistical analyses addressing the first research question will be conducted using the working sample of audiotaped tasks which will contain 104 recorded performances (26 students performing 4 tasks each). The working sample will result after subtracting the sub-sample to be used for rater training by randomly picking 6 examinees (with the corresponding 24 task performances) from the total of 32 examinees (with corresponding total of 128 task performances)

4.5.1 Statistical Analysis

The study will entail a Repeated Measures Analysis of Variance with a completely within design. The analysis will provide information about cued versus uncued conditions; dialogue versus monologue conditions, and the interaction between the cued/uncued and the dialogue/monologue conditions. This statistical analysis will be conducted on the sample of 26 examinees with the corresponding 104 audiotaped speech samples.

4.5.2 Discourse Analysis

In the second phase of the analysis the researcher will perform discourse analysis of the 24 transcribed speech samples from the six

case studies. Profiles of high and low fluency speakers will then be interfaced with four task-types in order to examine whether there are qualitative and quantitative differences in the discourse produced, and if so, whether they correlate with high or low fluency ratings.

4.6 Anticipated Outcomes

A. It is expected that task structuring involving manipulation of cued/uncued and dialogue/monologue conditions will have an impact on the amount of formulaic speech produced and the fluency ratings obtained on the tasks. More specifically, it is hypothesized that the cued tasks will generate more formulaic speech and that performance on them will be rated as more fluent than on the uncued tasks. It is expected that the uncued role-play will generate more negotiation of meaning and dispute for turn-taking than the cued role-play, with those more in control of the conversation being rated higher in fluency.

B. It is further hypothesized that the monologue tasks will lend themselves to more nonnativelike pause and hesitation phenomena and to a clearer picture of the speaker's interlanguage level than the dialogue tasks because the speakers will be required to display fluency at the discourse level rather than only at the utterance level. This will result in a number of possible outcomes: a) speech rate will be slower which may be rated as less fluent performance; b) overt monitoring and repair phenomena will be more conspicuous which might lead to lower fluency ratings; c) speech rate may remain relatively unaltered and fluency ratings will lie more heavily on linguistic competence (i.e., grammaticallity) in which case the presence or absence of effective strategic competence will respectively account for higher or lower fluency ratings, or d) the amount of formulaic speech will positively correlate with fluency ratings.

C. It is anticipated that formulaic speech will both quantitatively and qualitatively vary in the discourse produced to perform the four speaking tasks both within a given speaker's discourse across the fours tasks and for the aggregate group of speakers' discourse across the four tasks. In qualitative terms, the amount of formulaic speech produced in the performance of a given task is expected to positively correlate with fluency ratings so that the greater the amount of formulaic speech produced, the higher the speaker's fluency rating on a given task. In qualitative terms, it is expected that the higher the speaker's fluency level (as manifested by obtained fluency rating), the larger the chunks of formulaic speech s/he will produce and the greater the presence of the chaining strategy (instead of the grammatical integration strategy) on interclause boundaries.

APPENDIX A

Cued Role-Play

Purpose: Requesting a Service

Prompt: Both examinee and examiners get each a card with turn by turn directions of what each is expected to say in English.

Roles: Examinee acts as self; examiner acts as American language teacher who works at the school and teaches privately as well.

Situation: Examinee will need to give a presentation in English to the company's General Manager who will be coming from the US to visit the Brazilian branch. The examinee is feeling apprehensive about the presentation and decides to get some help. For such, s/he approaches the teacher to make arrangements for some private classes.

Role Play Cards for Cued Dialogue:

Examinee:

You work at a multinational company in Sao Paulo. You will need to give a presentation in English to the company's General Manager who will be coming from the US to visit the Brazilian branch. You are feeling apprehensive about the presentation, so you decide to get some help. For such, you will approach one of the school teachers to inquire about private classes.

Follow the cues:

1st- Greet teacher and tell him/her you need to talk to her.

2nd- Explain what the problem is. Answer whatever questions s/he may have.

3rd- Inquire about teacher availability as well as price, duration and site of private classes.

4th- Ask for an estimate of the frequency and duration of classes needed for you to be prepared for your presentation.

5th- Take into acount the expenses and time involved and come to a decision.

Teacher

The examinee works at a multinational company in Sao Paulo. S/he will need to give a presentation in English to the company's General Manager who will be coming from the US to visit the Brazilian branch. S/he is feeling apprehensive about the presentation, so s/he decides to get some help. Since s/he knows you teach privately, s/he approaches to inquire about private classes.

Procedures:

1. Let the examinee start the interaction. Throughout the roleplay give examinee the opportunity to take initiative and to speak uninterruptedly. However, feel free to ask questions should you need more information about the presentation or should something have remained unclear.

2. Be prepared to answer his/her questions about the private classes regarding price, location, and duration of classes.

3. When requested, give examinee an estimate of the number and frequency of classes necessary to prepare sim/her for the presentation. Try to make the examinee feel free to come to a decision about the classes and the arrangements without putting pressure on him/her.

Uncued role-play with examiner:

Roles: The examinee acts as self; the examiner (a teacher) assumes role of a foreigner currently in Brazil on business.

Situation: At a multinational company/office/job site. Ready to leave after a day's work.

Purpose/Function: Brazilian will try to convince foreigner to go out for a drink or a light dinner. The foreigner feels surprised by the unexpected invitation (which is typical of Brazilian spontaneity) and is somewhat reluctant.

Uncued Monologue: Personal narrative task:

Prompt:

An important holiday (E.g.: Christmas, The Jewish New Year, etc) is approaching, and you've decided to invite an American friend to spend it with your family since s/he is going to be away from his/her family. Your friend arrives punctually as customary in the US but of course nobody else has arrived yet (which is typical of Brazil). To entertain him/her you start talking about one of the relatives s/he will meet.

"Describe an interesting member of your family: one of your parents or grandparents, a brother or sister, cousin, aunt, or uncle. Be specific in describing the special characteristics that you think make him or her an interesting person. You may wish to give examples of things this person has said or done which illustrates these characteristics." (Stansfield, 1991)

Picture-cued Monologue:

Examinee will be given some time to study a picture sequence (series of pictures or line drawings depicting a story) and then s/he will be asked to narrate story in the past/hypothetical past.

REFERENCES

ACTFL (1986). ACTFL proficiency guidelines. NY: The American Council for the Teachers of Foreign Languages.

Austin, J. L. (1962) How to do things with words. Oxford: Clarendon.

Bachman, L. F. (1990). Fundamental considerations in language testing. NY: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1982). The construct validation of some components of communicative proficiency. TESOL QUARTERLY, 16(4), 449-464.

Bachman, L. F., & Savignon, S. J. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL Oral Interview. Modern Language Journal, 70 (4), 380-390.

Brown, J. D. (1988). Understanding research in second language learning: A teacher's guide to statistics and research design. New York: Cambridge University Press.

Brown, G., & Yule, G. (1983). Teaching the spoken language: An approach based on the analysis of conversational English. New York: Cambridge University Press.

Butler-Wall, B. A. (1986). The frequency of disfluencies in native and non-native conversational discourse. Doctoral Dissertation. University of California, Los Angeles.

Bygate, M. (1987). Speaking. Oxford: Oxford University Press.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language testing. Applied Linguistics, 1(1), 1-47.

Cathcart, R. (1986). Situational differences and the sampling of young children's school language. In Day, R. R. (Ed) Talking to learn: Conversation in second language acquisition. Rowley, MA: Newbury House. pp 119-140

Clark, J. L. D. (1986). Development of a tape-mediated ACTFL/ILR Scale based test of Chinese speaking proficiency. In Stansfield C. W. (Ed). Technology in language testing. Washington, D.C.: TESOL Publications. pp 130-145.

Clark, J. L. D., & Swinton, S. S. (1979). An exploration of speaking proficiency measures in the TOEFL context. TOEFL Research Reports No. 4. Princeton, NJ: Educational Testing Service.

Corder, S. P. (1973). Introducing Applied Linguistics. Hammonsworth: Penguin.

Doughty, C., & Pica, T. (1986). 'Information gaps': Do they facilitate second language acquisition? TESOL Quarterly, 20(2)

Duff, P. (1986). Another look at interlanguage talk: Taking task to task. In Day, R. R. (Ed) Talking to learn: Conversation in second language acquisition. Rowley, MA: Newbury House.

Ervin, G. L. (1983). (Ed). Pathways to language testing. Modern Language Journal, 67(4).

Finocchiaro, M. & Brumfit, C. (1983). The functional-notional approach: From theory to practice. New York: Oxford University Press.

Goffman, E. (1981). Forms of talk. Philadelphia: University of PA Press.

Graham, S. L. (1978). Using the FSI interview as a diagnostic evaluation instrument. In Clark, J. L. D. (Ed.). Direct testing of speaking proficiency: Theory and application. Princeton, NJ: Educational Testing Service.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. Language Learning, 24(2), 287-298.

Harlow, L. L. (1990). Do they mean what they say? Sociopragmatic competence and second language learners. Modern Language Journal, 74 (3), 328-351.

Hatch, E. V. (1978).(Ed) Second language acquisition: A book of readings. Rowley, MA: Newbury House.

Hatch, E. V. (1978). Discourse analysis and second language acquisition. In Hatch, E. V. (Ed) Second language acquisition: A book of readings. Rowley, MA: Newbury House.

Herrmann, F. (1990). Instrumental and agentive uses of the computer: Their role in learning French as a forein language. Doctoral Dissertation. Stanford University.

Hieke, A. E. (1984a). Linking as a marker of fluent speech. Language and Speech,27 (4), 343-354.

Hieke, A. E. (1984b). Oral fluency and its evaluation. ERIC Document Reproduction Service No. ED 250 910.

Hieke, A. E. (1988). Spoken language phonotactics. ERIC Document Reproduction Service No. ED 290 320.

Higgs,T. V. (Ed) (1982). Curriculum, competence, and the foreign language teacher. Skolkie, Ill: National Textbook Co.

Higgs, T. V. (1984). (Ed). Teaching for proficiency: The organizing principle. Lincolnwood, Il: National Textbook.

Higgs, T. V., & Clifford, T. (1982). The push toward communication. In Higgs,T. V. (Ed) (1982). Curriculum, competence, and the foreign language teacher. Skolkie, Ill: National Textbook Co.

Hymes, D. (1972). On communicative competence. In J. B. Price & J. Holmes. (Eds). Sociolinguistics. Harmondsworth: Penguin, pp. 269-293.

James, C. J. (Ed). (1985). Foreign language proficiency in the classroom and beyond. Lincolnwood, Il: National Textbook.

Jefferson, G. (1979). Error correction as an interactional resource. Language and Society, 3, 181-199.

Jull, D. (1987). Traditional techniques in the teaching of pronunciation. TESL-Talk, 17 (1), 118-124.

Kramsch, C. J. (1981). Discourse analysis and second language teaching. Language in Education no. 37. Washington, DC: Center for Applied Linguistics.

Kramsch, C. (1986). From language proficiency to interactional competence. Modern Lnguage Journal, 70 (4), 366-372.

Krashen, S. D., & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. Language Learning, 28, 283-300.

Labov, W. (1972). Sociolinguistic patterns. Philadelphia, PA: University of Pennsylvania Press.

Langer, J. A. (1987). A sociocognitive perspective on literacy. In Langer, J. A. (Ed). Language. literacy and culture: Issues of society and schooling. Norwood, NJ: Ablex.

Langer, J. A. et al. (1990). Meaning construction in school literacy tasks: A study of bilingual students. American Educational Research Journal, 27 (3), 427-471.

Long, M. H. (1980). Input, interaction, and second language acquisition. Ph.D dissertattion. UCLA.

Long, M. H. (1981). Input, interaction, and second language acquisition. In Winitz, H. (Ed). Native language and foreign language annals of the New York Academy of Sciences, 379, 259-278.

Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. Studies in Second Language Acquisition, 5 (2), 177-194.

Lowe, P. Jr. (1988)

Mullen, K. A. (1978). Direct evaluation of second language

proficiency: The effect of rater and scale in oral interviews. Language Learning, 28 (2), 301-308.

Murphy, D. (1986). Communication and correction in the classroom. ELT Journal, 40 (2), 146-151.

Nattinger, J. R. (1985). Lexical phrases, functions and vocabulary acquisition. ERIC Document Reproduction Service No. 268 823.

Nunan, D. (1989). Designing tasks for the communicative classroom Cambridge: Cambridge University Press.

Oller, J. W. Jr. (1979). Language tests at school. New York: Longman.

Oller, J. W. Jr. (1983a). A consensus for the eighties? In Oller 1983b, 351-356

Oller, J. W. Jr. (Ed). (1983b). Issues in language testing research. Rowley, MA: Newbury House.

Olynyk, M., Sankoff, D., & D'Anglejan, A. (1983). Second language fluency and the subjective evaluation of officer cadets in a military college. Studies in second Language Acquisition, 5(2), 213-249.

Omaggio, A. C. (1986). Teaching language in context. Boston, MA: Heinle & Heinle.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and native fluency. In J. C. Richards & Schmidt, R. W. (Eds). Language and communication. New York: Longman, pp 191-226.

Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts? Language, 53, 560-573.

Porter, P. (1983). How learners talk to each other: Input and interaction in task-centered discussions. Paper presented at the 1983 TESOL Convention. Toronto.

Raupach, M. (1980). Temporal variables in first and second language speech production. In Dechert, H. W. & M. Raupach (Eds). Temporal variables in speech: Studies in honor of Frieda Goldman-Eisler. The Hague: Mouton.

Raupach, M. (1984). Formulae in second language speech production In Dechert, H. W., Mohle. D., & Raupach, M. (Eds). Second language productions.

Richard-Amato, P. A. (1987). Making it happen. White Plains, NY: Longman.

Richards, J. C., & Schmidt, R. W. (Eds). (1983). Language and

communication. New York, NY: Longman.

Riggenbach, H. R. (1989). Nonnative fluency in dialogue versus monologue speech: A microanalytic approach. Doctoral Dissertation. University of California, Los Angeles.

Samuda, V. (1990). Task as a learner construct. Paper presented at the 1990 Annual TESOL Convention. San Francisco, CA.

Schachter, J. (1986a) In search of systematicity in interlanguage production. Studies in Second Language Acquisition, 8, 119-133.

Schachter, J. (1986b). Three approaches to the study of input. Language Learning, 36(2), 211-226.

Schaefer, C. F. (1980). Episodic memory, semantic memory, and fluency. International Review of Applied Linguistics in Language Teaching, 18 (4), 321-325.

Schonfield, P. (1987). Communication strategies -- The researcher outmanoeuvred? Applied Linguistics, 8 (3), 219-232.

Selinker, L. (1972). Interlanguage. International Review of Applied Linguistics, 10, 209-231.

Stansfield, C. W. (1986). Towards communicative competence testing: proceedings of the second TOEFL Invitational Conference, TOEFL Research Report 21, Princeton, NJ: Educational Testing Service.

Underhill, N. (1987). Testing spoken language. Cambridge: Cambridge University Press.

Vallette, R. M. (1977). Modern language testing. Second edition. NY: Harcourt Brace Jovanivitch.

Wong-Fillmore, L. (1976). The second time around. Doctoral dissertation. Stanford University.

Young, R. (1984). Negotiation of outcome and negotiation of meaning. TESOL QUARTERLY, 18(3), 525-526.

Yorozuya, R. (1982). Evaluation of oral fluency in English as a Foreign Language. Master's Thesis. Hiroshima University, Japan. ERIC Document Reproduction Service No. ED 230 018.

6. APPENDICES: **Appendix A**

Letter of Consent from students (in Portuguese)

Nome do Aluno: _____

Querido aluno(a):

Meu nome e Roseli Ejzenberg. Sou uma brasileira fazendo o curso de doutorado nos Estados Unidos na State University of New York at Albany. Tenho sido colega profissional do pessoal da Seven desde os tempos em que trabalhavamos como professores do Alumni (onde depois fui supervisora de cursos avancados) e mesmo agora ja residindo ha 4 anos no exterior temos mantido contato profissional e pessoal.

Em marco de 1991 representei a escola no Congresso Internacional sobre o Ensino de Ingles para falantes de outros idiomas (1991 Annual TESOL Conventional) no qual o trabalho da equipe da escola com os Placement Exams foi elogiado por conhecidos especialistas na area de avaliacao e testes. Agora pretendemos dar continuidade ao trabalho realizando um estudo para aprender mais sobre os testes orais de alunos de cursos avancados, o que e o tema da minha dissertacao.

Os resultados deste trabalho deverao repercutir nao somente em beneficio da Seven e de voce como tambem serao uma importante contribuicao para esta area de pesquisa. Para a realizacao do estudo e muito importante contar com sua colaboracao comparecendo ao teste oral a ser realizado com um/uma dos/das professores/porfessoras da escola num horario fora de aula que seja conveniente para voce. O teste nao dura mais que 35 minutos e sera gravado com um gravador ja que eu nao estarei presente para nao inibi-lo.

Na segunda fase do estudo eu prepararei 6 fitas com pedacos dos exames de cada aluno misturados de tal maneira que nao se saiba a identidade do examinando. Dai tres professores da escola vao escutar as fitas e dar nota. Podem ficar tranquilos que as notas nao serao usadas como avaliacao dentro dos cursos nem divulgadas. Finalmente, quando eu voltar para os Estados Unidos eu vou fazer uma analise linguistica dos testes e analisar os dados estatisticamente para chegar a conclusao de como se deve estruturar os exames orais de cursos avancados como o seu.

Por favor indique tambem abaixo se voce teria outra meia hora para conversar comigo sobre seu feedback a respeito do teste. Este papo pode ser por telefone ou pessoalmente. Contando com seu apoio.

Roseli Ejzenberg

Concordo com que meu teste seja gravado. Sim __ Nao __
Concordo em ser "entrevistado". Sim __ Nao __
Se for o caso indique preferencia: Pessoalmente __ Por fone __

_____
Assinatura do(a) Aluno(a)

Letter of Consent from students (Translation)

Dear Student:

My name is Roseli Ejzenberg. I am a Brazilian doctoral student at State University of New York at Albany. The school owners and teachers and I have been colleagues ever since the time when we worked together as teachers at the Alumni Association where I also worked as supervisor of Advanced Courses. We've maintained professional and personal contact even though I've been in the US for 4 years now.

In march of 1991 I represented your school at the 1991 Annual TESOL convention where the work of Seven Language & Culture on Placement Exams was praised by known experts in the area of testing. Now, we intend to continue our cooperation by conducting a study on oral testing of the more advanced students, which is the topic of my dissertation.

The results of this study should benefit not only your school but alsc the field of reserach. For such, it is very important to count on your cooperation in the form of your attending a testing session to be conducted outside of class with one of the school teachers at a time of your convenience. The test won't last longer than 35 minutes and will be audiotaped. I won't be present to avoid inhibiting you.

In the second phase of the study I'll prepare 6 tapes with parts of the exam scrambled so that the examinee's identity can't be disclosed. Then 3 of the school teachers will listen to the tapes and rate oral performance. Rest assured that your grades won't be computed into your course grades or divulged in any form. Finally, after I go back to the States I'll conduct a linguistic analysis and study the data statistically in order to get to a conclusion about how we should test advanced students of foreign languages like yourself.

Please also indicate below if you would be willing to talk to me personally or by phone to give me some feedback about the test. I count on your cooperation.

Roseli Ejzenberg

I give my consent for my test to be audiotaped. Yes __ No __
I am willing to be interviewed about the test. Yes __ No __
If appropriate indicate preference: Personnally __ By phone __

_____
Student's signature

APPENDIX B

Cued Dialogue

Purpose: Requesting a Service

Prompt: Both examinee and examiners get each a card with turn by turn directions of what each is expected to say in English.

Roles: Examinee acts as self; examiner acts as American language teacher who works at the school and teaches privately as well.

Situation: Examinee will need to give a presentation in English to the company's General Manager who will be coming from the US to visit the Brazilian branch. The examinee is feeling apprehensive about the presentation and decides to get some help. For such, s/he approaches the teacher to make arrangements for some private classes.

Role Play Cards for Cued Dialogue:

Examinee:

You work at a multinational company in Sao Paulo. You will need to give a presentation in English to the company's General Manager who will be coming from the US to visit the Brazilian branch. You are feeling apprehensive about the presentation, so you decide to get some help. For such, you will approach one of the school teachers to inquire about private classes.

Follow the cues:

1st- Greet teacher and tell him/her you need to talk to her.

2nd- Explain what the problem is. Answer whatever questions s/he may have.

3rd- Inquire about teacher availability as well as price, duration and site of private classes.

4th- Ask for an estimate of the frequency and duration of classes needed for you to be prepared for your presentation.

5th- Take into acount the expenses and time involved and come to a decision.

Teacher

The examinee works at a multinational company in Sao Paulo. S/he will need to give a presentation in English to the company's General Manager who will be coming from the US to visit the Brazilian branch. S/he is feeling apprehensive about the presentation, so s/he decides to get some help. Since s/he knows you teach privately, s/he approaches to inquire about private

classes.

Procedures:

1. Let the examinee start the interaction. Throughout the roleplay give examinee the opportunity to take initiative and to speak uninterruptedly. However, feel free to ask questions should you need more information about the presentation or should something have remained unclear.

2. Be prepared to answer his/her questions about the private classes regarding price, location, and duration of classes.

3. When requested, give examinee an estimate of the number and frequency of classes necessary to prepare sim/her for the presentation. Try to make the examinee feel free to come to a decision about the classes and the arrangements without putting pressure on him/her.

Uncued dialogue with examiner:

Roles: The examinee acts as self; the examiner (a teacher) assumes role of a foreigner currently in Brazil on business.

Situation: At a multinational company/office/job site. Ready to leave after a day's work.

Purpose/Function: Brazilian will try to convince foreigner to go out for a drink or a light dinner. The foreigner feels surprised by the unexpected invitation (which is typical of Brazilian spontaneity) and is somewhat reluctant.

Uncued Monologue: Personal narrative task:

Prompt:

An important holiday (E.g.: Christmas, The Jewish New Year, etc) is approaching, and you've decided to invite an American friend to spend it with your family since s/he is going to be away from his/her family. Your friend arrives punctually as customary in the US but of course nobody else has arrived yet (which is typical of Brazil). To entertain him/her you start talking about one of the relatives s/he will meet.

"Describe an interesting member of your family: one of your parents or grandparents, a brother or sister, cousin, aunt, or uncle. Be specific in describing the special characteristics that you think make him or her an interesting person. You may wish to give examples of things this person has said or done which illustrates these characteristics." (Stansfield, 1991)

Picture-cued Monologue:

Examinee will be given some time to study a picture sequence (series of pictures or line drawings depicting a story) and then

s/he will be asked to narrate story in the past/hypothetical past. (Picture sequence enclosed.)

Appendic D: RATING FORM

Rating Session No. _____ Rater: _____

Dear Rater: Circle one of the numbers ranging from 1 (very low fluency) to 6 (very high fluency) on the rating scale to indicate your rating of the examinee's fluency at performing the task. The tape will be paused for one minute after you listen to each task for you to jot down any comments on why you gave the examinee a certain rating. This procedure is optional for ratings ranging from 2 to 5 and required for ratings of 1 or 6. Please be careful about recording each rating for each examinee in the sequence that appears on the tape. If you make a mistake or change your mind, just erase the previous rating and/or commemt and write the new information in the same space alloted for that examinee.

Examinee No. 1 FLUENCY RATING

Very Low fluency					Very High fluency
1	2	3	4	5	6

Comments: _____

_____

_____

Examinee No. 2 FLUENCY RATING

Very Low fluency					Very High fluency
1	2	3	4	5	6

Comments: _____

_____

_____

Examinee No. 3 FLUENCY RATING

Very Low fluency					Very High fluency
1	2	3	4	5	6

Comments: _____

_____

_____

Examinee No. 4 FLUENCY RATING

Very Low fluency — Very High fluency

1 2 3 4 5 6

Comments: _____

_____

_____

Examinee No. 5 FLUENCY RATING

Very Low fluency — Very High fluency

1 2 3 4 5 6

Comments: _____

_____

_____

Examinee No. 6 FLUENCY RATING

Very Low fluency — Very High fluency

1 2 3 4 5 6

Comments: _____

_____

_____

Examinee No. 7 FLUENCY RATING

Very Low fluency — Very High fluency

1 2 3 4 5 6

Comments: _____

_____

_____

Appendix E **SEQUENCING**

Although there is an effect of sequencing on performance it is most crucial when comparing the first and the last task performed, i.e., when the examinee had no previous task practice versus when the examinee had three previous task performances. If we combine this hypothesis with the one about the effect of performing the cued or uncued dialogue first and with our counterbalancing dialogue/monologue in the two halves of each task (i.e., there is one dialogue and one monologue in each half) we would end up with the following scenario:

To make sure all tasks appear at least once in first position in the sequencing of tasks we would select from the table above three tasks from the UD-first condition and three from the CD-first condition (Degrees of freedom= 4 - 1= 3) as it is shown in the table below.

Table 2. Test Sequencing:

TASK SEQ.	**"UD-first" test**	**"CD-first" test**
No.	? ? ?	? ? ?
1st	**UM CM UD**	**CD UM CM**
2nd	? ? ?	? ? ?
3rd	? ? ?	? ? ?
4th	? ? ?	? ? ?
	Subtotal: 3 combinations	3 combinations

Total: 6 possible combinations

However, we also need to make sure all tasks have a chance to appear at least once on fourth position. My solution to this was to pick 6 among all possible combinations (16) derived in Table 1 that would meet the following criteria:

a) Half from UD-first test (N=3); half from CD-first test (N=3)

b) Each task will be in first position and in last position at least once.

c)One test from each category (i.e., UD-first versus CD-first) would have the UD and CD tasks in adjacent positions so that one would directly precede or follow the other. E.g: Test No. 3 and Test No. 11.

d)The positioning of the monologue tasks on one extreme of the sequence, as the first task of the test, was deliberately manipulated since usually it is taken for granted that dialogue tasks should occupy that slot instead. Furthermore, because the CM (the picture-

cued task) is the one with the least amount of face validity and authenticity because (...) "the content of the set of pictures is usually known to the hearer before the speaker tells the story (and the speaker knows this), which rather impairs the genuineness of the communication." (Schonfield, 1987, p. 219) it seemed potentially interesting to investigate whether the positioning of the CM (picture-cued task) as first task would throw off the examinees' performance in any way.

This lengthy screening resulted in the preliminary selection of the following tasks for the study:

Table 3- TASK SEQUENCING: TEST SELECTION

TASK SEQ.	**"UD-first" test**			**"CD-first" test**		
No.	**3**	**5**	**8**	**10**	**11**	**13**
1st	UM	CM	UD	CD	UM	CM
2nd	UD	UD	CM	UM	CD	CD
3rd	CD	UM	CD	CM	UD	UM
4th	CM	CD	UM	UD	CM	UD
	Subtotal: 3 combinations			3 combinations		
	Total: 6 test formats					